Journal of Responsible Technology 19 (2024) 100089

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.sciencedirect.com/journal/journal-of-responsible-technology

Decoding faces: Misalignments of gender identification in automated systems

Elena Beretta [a,*], Cristina Voto [b], Elena Rozera [c]

[a] Vrije Universiteit Amsterdam, Faculty of Science, Department of Computer Science, User-Centric Data Science group, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands
[b] University of Turin, ERC FACETS, Department of Philosophy and Educational Sciences, via Sant'Ottavio, 20, 10124 Torino TO, Italy
[c] Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands

ARTICLE INFO

Keywords:
Automatic gender recognition
Gender identity
Face recognition

ABSTRACT

Automated Facial Analysis technologies, predominantly used for facial detection and recognition, have garnered significant attention in recent years. Although these technologies have seen advancements and widespread adoption, biases embedded within systems have raised ethical concerns. This research aims to delve into the disparities of Automatic Gender Recognition systems (AGRs), particularly their oversimplification of gender identities through a binary lens. Such a reductionist perspective is known to marginalize and misgender individuals. This study set out to investigate the alignment of an individual's gender identity and its expression through the face with societal norms, and the perceived difference between misgendering experiences from machines versus humans. Insights were gathered through an online survey, utilizing an AGR system to simulate misgendering experiences. The overarching goal is to shed light on gender identity nuances and guide the creation of more ethically responsible and inclusive facial recognition software.

1. Introduction

In the rapidly evolving landscape of Artificial Intelligence, where the interaction between technology and human identity is increasingly scrutinized, Automated Facial Analysis (AFA) emerges as a critical domain for ethical and societal reflection. Employing advanced deep neural networks the discipline is predominantly composed of two key processes: facial detection and facial recognition. Facial detection pertains to the task of identifying the existence of a face within a digital image or a video stream. Upon successful detection, facial recognition is undertaken to distinguish specific individuals based on their unique facial attributes (Scheuerman et al., 2019). The reliability and precision of these systems have seen remarkable advancements over time. This rapid growth has led to their widespread adoption across a diverse range of sectors. Initially, facial recognition technology, like Automated Facial Analysis, has been primarily used for security (Balla & Jadhao, 2018; Karovaliya et al., 2015) and law enforcement purposes (Bradford et al., 2020; Kaur et al., 2020). However, its applications now extend far beyond these traditional domains. In the realm of recruitment, facial recognition is being explored to streamline hiring processes and assess candidate suitability (Majumder & Bhattacharya, 2021; Mujtaba & Mahapatra, 2019). Business applications are also emerging, with companies leveraging this technology for customer engagement and personalized marketing (Christopher Hlongwane et al., 2021; Zeng & Chiu, 2021). In education, it is used for monitoring student engagement and attendance (Andrejevic & Selwyn, 2020; Krithika et al., 2017), while the healthcare sector is exploring its use in patient identification and diagnosis (Bisogni et al., 2022). Furthermore, the analysis of facial expressions (Mane & Shah, 2019; Tian et al., 2005) and emotions (Wolf, 2015) through facial recognition is gaining traction, providing valuable insights in psychological and behavioral studies. However, the integration of these advanced technologies into the fabric of our society necessitates a careful and thorough consideration of the ethical implications that accompany their use. The expansive use of these systems in diverse societal contexts can lead to cultural misunderstandings and misrepresentations. For example, the way these systems interpret and categorize facial features can be heavily influenced by the cultural biases inherent in their programming and data sets. This can result in a technology that, albeit inadvertently, reinforces stereotypical or culturally insensitive portrayals of certain groups (Buolamwini & Gebru, 2018). Hence, despite advancements in their accuracy, these systems are not immune to biases, which can result in discriminatory practices. The

* Corresponding author.
E-mail address: elena.beretta@vu.nl (E. Beretta).

https://doi.org/10.1016/j.jrt.2024.100089

Available online 17 June 2024

impact of these biases is not merely theoretical but has been documented in various studies and incidents over the years, spotlighting discrepancies across diverse demographic subsets. A stark example of this issue was the 2015 incident where Google Photos erroneously tagged two black individuals as gorillas (Barr, 2015), highlighting the severe consequences of these biases. Further research has deepened our understanding of these issues, indicating that commercial gender classification algorithms' performances can be significantly influenced by skin color (Williford et al., 2020), often demonstrating superior performance for lighter-skinned males and remarkably inferior performance for darker-skinned females (Buolamwini & Gebru, 2018). More recently, in 2019, a study probing face detection rates using an array of models identified a substantial bias towards particular ethnicities. This study revealed significant differences in the impostor and genuine distributions between African-American and Caucasian cohorts, with the former exhibiting a higher false match rate and a lower false non-match rate (S et al., 2019). These biases are not just technological failures but also reflect the cultural and societal narratives embedded within these technologies. Cumulatively, these studies underscore the exigent necessity for constant refinement and rigorous evaluation of facial recognition systems to rectify these ingrained biases, thus ensuring the technology's equitable implementation across all demographic strata. This need for refinement is particularly evident in the domain of AFA, specifically within Automatic Gender Recognition (AGR), where machine learning algorithms are employed for the computational categorization of individuals' gender from photographs or videos. These AGR systems, utilized in a wide range of applications from human-computer interaction to commercial development, demographic research, and entertainment, employ physical markers such as the structure of lips, eyes, and cheeks for gender predictions. However, recent empirical investigations have unveiled considerable disparities in these systems. AGR predominantly approach gender classification through binary lens, categorizing individuals as either male or female. This oversimplified perspective, lamentably, overlooks non-binary or fluid gender identities, inadvertently perpetuating marginalization and eliciting profound societal implications, especially for people who identify as transgender and/or non-binary. Gender identity, as a result of a complex interplay of factors, evolves through continuous social embodiments and representations. In this context, the dissemination of concepts such as intrinsic sex and enduring masculinity or femininity functions as a pivotal component within a strategic framework, dedicated to the ideological support of the binary structure of identity. Nonetheless, directing attention to the performative nature of identity serves as a reminder that gender is also molded through performance, emphasizing that, for this reason, it remains a dynamic and non-fixed dimension. Moreover, misgendering — the act of addressing someone in a way that does not align with their self-identified gender— has been linked to significant psychological impacts, including harm to selfesteem, mental health deterioration, and heightened social stigma (Dembroff & Wodak, 2018; Keyes, 2018; Sue, 2010). This underscores the critical importance of developing AGR systems that are not only technically proficient but also deeply attuned to the complexities of human identity, respecting the individual's self-perception and societal representation.

In recognizing the potential repercussions of the lack of inclusion in Automatic Gender Recognition (AGR) systems, our research underscores the imperative for a holistic approach that extends beyond technical solutions. This approach demands careful consideration of the ethical and societal dimensions that are integral to the development and deployment of these technologies. Our investigation, therefore, contributes to the broader academic dialogue by examining the societal impacts of facial recognition technology, with a particular focus on the complexities of gender identity. By addressing the profound implications of misgendering by automated systems, we hope to shed light on the complexities of identity that are often overlooked in conventional technological frameworks. We emphasize the significance of facial images as not just mere visual markers, but as powerful signifiers that intertwine with deeper societal narratives on identity. It's worth noting that this constitutes an initial version of our study, and we anticipate subsequent iterations and expansions in our forthcoming releases. This initial study, thus, serves as a preliminary exploration of research that centers on the issue of recognizing identity and, consequently, otherness, in the age of machines.

1.1. Problem statement

The modus operandi by which individuals conceptualize and explicate their gender identity is disparate from the outward presentations that society employs for gender categorization. If, as human beings, we engage with the experience of gender identity through a series of performances to embody a specific identity rather than another, the social representations of gender identity tend to attain such levels of recognizability that the aesthetic experience of gender can be readily identified, mediated, and communicated. The significant result of these representations is evident today in the training of recognition technologies to quantify the characteristics by which identities can be identified, especially in relation to computer vision (AFA, AGR, etc.). The predilection of state-of-the-art automated facial recognition technologies to anchor on physical characteristics, which do not consistently align with an individual gender identity, creates a gap that acts as a source of complications within these systems. It thus becomes imperative to understand gender identity beyond mere visual attributes, in order to cultivate gender recognition software that are inclusive. The aim of this study is to undertake an examination of the emergent issue concerning the intra-action (Barad 2007) between facial detection technology and non-binary gender identification, conducting a multidisciplinary research and integrating insights and methods from both qualitative and quantitative approaches in the analytical and research methodology. Identity is consistently shaped by a series of interactions, implications, responses to stimuli, and prompts within the sociocultural dimension. Acknowledging the concept of identity as a performative experience, we recognize the dialogical and open dimension wherein our identity is in constant dialogue and redefinition with an otherness through which it interacts via identification. This alterity may take the form of human, non-human animal, or technological otherness As stated by Karen Barad:

> *Since individually determinate entities do not exist, measurements do not entail an interaction between separate entities; rather, determinate entities emerge from their intra-action. I introduce the term "intra-action" in recognition of their ontological inseparability, in contrast to the usual "interaction", which relies on a metaphysics of individualism (in particular, the prior existence of separately determinate entities) (2007, 128).*

Following Barad, the impetus for this investigation emanates from a broadly encompassing query: how do Automated Facial Analysis and Automatic Gender Recognition technologies intra-act on our society and the manner in which we construct our identities? Approaching from an alternative perspective, how does the binary constructivism inherent in facial recognition algorithms impact on an individual's self-perception of identity?

1.2. Research questions

In light of the intricate relationship between automated facial recognition technologies and the spectrum of gender identification, we have delineated a set of research questions to steer our investigation:

RQ1: To what extent does an individual's gender identity align with societal norms, and how is this reflected in their appearance, particularly their facial expression?
RQ2: Is there a perceived difference between being misgendered by a machine and being misgendered by a human?

These research inquiries, rooted in our overarching investigation into the societal implications of facial recognition in the context of individual identity formation, endeavor to illuminate the intricacies of gender identity that extend beyond superficial visual attributes. Furthermore, we seek to delve into the consequences of individuals encountering misgendering by automated systems. It is prudent, in this research endeavor, to consider facial images as artifacts that not only capture visual characteristics but also encapsulate the broader discourses surrounding identity. Consequently, the data extracted from these images emerges as a repository of discussions and representations related to identity. Our primary objective is to unearth insights that can inform the development of more inclusive facial recognition software solutions.

Our focus lies in contemplating how human-machine interactions can impact the perception of gender identity and shape identity performances and experiences. In this context, we are keen on addressing the issue from both the perspectives of performativity and the biases that technologies may perpetuate. Crucial for achieving our objectives will be to ground our analysis in a cross-disciplinary perspective—a lens through which we can facilitate a dialogue between challenges in computer science, particularly in computer vision, and issues in the humanities, specifically in the semiotics of post-human identity.

To realize this objective, we designed a survey to be administered through an online web application. This survey not only offers an interactive experience that leverages an AGR (Automated Gender Recognition) system to analyse instances of misgendering but also aims to examine the alignment of gender identity with societal norms and the implications of misgendering by automated systems. This approach serves as an initial experiment, laying the groundwork for a more comprehensive and in-depth exploration in the future.

2. Misgendering in contemporary intersections between AI and identity

Automated Facial Analysis (AFA) stands at the intersection of technological innovation and complex societal constructs, presenting both opportunities and challenges. The ethical dimensions of gender classification within AFA underscore the need for a nuanced understanding of how technology can both reinforce and challenge established gender norms. Delving deeper into the realm of identity, the performative nature of perception emerges, shaped by actions, interactions, and artefactual influences. Misidentification (see Section 2.2), whether by human judgment or algorithmic determinations, brings to the fore the profound implications of recognition and its absence. As individuals navigate spaces dominated by majority perceptions, strategies of adaptation and resistance come into play, highlighting the dynamic intra-action (Barad 2007) between individual identities and collective norms. This section seeks to provide a comprehensive overview, connecting digital and human aspects, as well as individual and collective experiences.

2.1. Gender classification and ethical implications in automated facial analysis

The study of biases and misgendering in Automatic Gender Recognition (AGR) systems has gained traction in recent years. Several studies have been conducted to understand the implications of these biases, especially on non-binary and transgender individuals. Above, we discussed misgendering, defined as the act of referring to someone in a manner inconsistent with their recognized gender identity. We also highlighted the current inadequacy of systems to accurately classify individuals who do not conform to the traditional gender binary (Buolamwini & Gebru, 2018; Scheuerman et al., 2019). In the same study, Buolamwini et Gebru, 2018, focused on the intersectionality of gender and race in AI system. They identified significant disparities in accuracy rates among these groups, showing, for example, how the False Positive Rate for women tends to be twice as high as the FPR for men, meaning that people are more likely to be classified as women than men. In a paper from 2018, Keyes (2018) presented an analysis on the studies in the field of AGR, attempting to comprehend how researchers operationalise gender, with a particular focus on transgender and gender non-conforming individuals. This analysis found that, despite the importance of the topic, most research focuses on binary gender classification, and that the inaccuracy of AGR tools might cause harm and social exclusion towards these individuals. A previous study by Kumar et al. (2016) on transgender individuals showed that being misgendered by an AGR is considered worse than being misgendered by human beings. The same concept has been investigated by Hamidi et al. (2018) who conducted a series of interviews with gender non-conforming individuals on their attitudes towards AGR. The majority of the participants perceived being misgendered by AGR systems to be more detrimental compared to being misgendered by humans. The participants attributed this perception to the belief that these systems introduced an additional layer of invalidation, potentially reinforcing and perpetuating traditional gender norms and amplifying the already adverse consequences of everyday misgendering. Furthermore, several participants expressed concerns that misgendering by an AGR system was particularly distressing due to the expected objectivity and precision of technology.

To address the issue of misgendering, researchers have commenced investigations into alternative methodologies. However, it is imperative to emphasize that even these approaches can be undermined by ethical issues and methodological imperfections arising from the absence of solicited feedback from individuals regarding the attributed *gender* label. This deficit of reciprocal engagement begets pronounced ethical apprehensions, as it disregards the volition and informed consent of individuals involved in the study. Biased databases wield a considerable influence over the disparities in predictive accuracy across a spectrum of tasks. Prevalently, benchmark databases are constructed through the initial utilization of facial detection algorithms to identify facial visages from online repositories (Huang et al., 2007). Nevertheless, the potential existence of systematic inaccuracies inherent within these facial detection algorithms stands to imbue benchmark databases with an inherent bias, particularly in cases characterized by a paucity of demographic diversity. In the year 2020, Wu and colleagues undertook an inquiry with the intent to augment an extant binary classification framework, encompassing within it a non-binary gender classification category. The classifier attained commendable accuracy rates in the domain of gender prediction within nonbinary populations, with the most proficient model achieving an accuracy score of 91.97 % (Merler et al., 2019; Ryu et al., 2017; Wu et al., 2020).

In the context of the ongoing dialogue between computer vision, particularly automated face recognition, and the humanities, with a focal point on inquiries into identity perception, we propose three pivotal terms, Identification, Misidentification, and Disidentification. Drawing from cultural and performative studies, we adapt these concepts to the technical domain of computer vision, offering a lens to examine how technology interacts with and impacts identities.

Identification in computer vision refers to the process where systems correctly match a detected face with an individual's unique identity. Misidentification, on the other na,d occurs when facial recognition systems erroneously match a detected face with an incorrect identity. The misidentificaion can stem not only from technical factors like varying lighting or angles but also from cultural biases embedded within the models. Such biases might cause systems to misinterpret facial features based on cultural backgrounds, leading to disproportionate rate of errors for certain demographic groups. Disidentification involves methods used to alter facial features to prevent or obstruct automated recognition. This can include the use of makeup, accessories, or other alterations aimed at changing the typical facial characteristics analyzed by these systems. Disidentification strategies are a form of resistance against surveillance and a means to maintain privacy and autonomy in

the face of pervasive facial recognition technologies.

2.2. Performing identities: navigating identification, misidentification, and disidentification

Gender identity evolves through an extensive learning process, where the experiences and behaviors of an individual interweave with those of others. It is culture, distinct from nature, that shapes the construction of sexual difference and discrimination (Fausto-Sterling & Stein, 2004) . The variance in sexual identity among individuals is socially and culturally constructed, distinct from the biological differences between sexes determined by nature. This cultural construction regulates the social division of roles and the processes of socialization between women and men. Being part of a gender primarily signifies belonging to a community.

Consequently, gender identity is increasingly acknowledged to operate along a continuum and its multifarious nature is such that a mere designation falls short of capturing its essence (De Lauretis, 1987; Preciado, 2013). Furthermore, gender expressions might be subject to variations within different societal settings (Beauvoir, 1948; Butler, 2002; Halberstam, 2019). In this survey, we will also dive into the concept of sociocultural gender performativity, specifically examining its relation to changes in the visual representation of the participants' gender identity and to the recognition in a human-machine intra-action. Identity is formed by how we deal with the different meanings and stories that society places on each of us (Hall & Du Gay, 1996) (Crenshaw, 2017).

Identity is always the result of the intersection of various aspects—such as those related to ethnic identity, class identity, age, and, of course, gender identity—through which an individual's identification is configured. A useful tool for framing the ways in which gender becomes identity and for analyzing how each intersection contributes to unique experiences of oppression and/or privilege, thinking in terms of intersectionality allows the recognition of the performative nature of identity. Identity is, in fact, always the result of plural and singular narratives that traverse society and, at the same time, the effect of inscriptions marked on the skin of each of us by discourses related to processes of identification. It is in this sense, we can think of identities not as fixed or essentialistic ontologies but as spaces, even of tension, where the right to self-determination can be exercised.

It is important to understand that identity is not something fixed or set in stone. It is a space where different parts of who we are, such as background, social class, race, age, abilities, and gender, all interact and create tension. Thinking of identity as a process that changes and evolves, shifts the focus from mere individualistic subjectivity to a dynamic interplay through which the individual engages in relations with others, fostering a sense of community and belonging. Identity is a complex interplay of the three, already mentioned, main concepts that lie at the foundations of this research, constituting the three macro-sections of our survey: identification, misidentification and disidentification (Butler, 2004). Each concept represents a different performative process that dynamically shapes our understanding of self and our interactions within the societal fabric.

2.2.1. Crafting self through action: the performative process of identification

Identification, by the psychoanalytic definition first developed by Sigmund Freud and later discussed by Jacques Lacan and others, constitutes an intricate process where an individual consciously or unconsciously assimilates characteristics of another, transforming their own self-identity in accordance to that model. This mechanism is fundamental to the development and evolution of our personalities and is also shaped by the engagements that, increasingly on a daily basis, we have with technologies. As we interact with different individuals in various social contexts, we continually integrate different aspects of their identities, reshaping our own (Lacan, 1962; Meissner, 1970). Furthermore, in contemporaneity, in the complex landscape of identity formation, the artifactual dimension emerges as a critical and often overlooked component. Identification processes are inextricably intertwined with the material dimension of society, where tangible artifacts such as photographs, identification documents, and technologies for automated recognition play a pivotal role (Belting, 2017). These artifacts not only serve as external markers of identity but also act as catalysts for the processes of self-recognition and identity construction (Sekula, 1986). They are not mere static representations but dynamic agents that influence our self-perception, guiding us in the assimilation of characteristics from the world around us. Consequently, delving into this artifactual dimension is essential for a comprehensive examination of the multifaceted nature of identification processes and their profound impact on the development and evolution of our individual and collective identities (Tian et al., 2005). In identity studies, *identification* denotes a performative process, extending beyond a mere act. It's an ongoing, dynamic effort to define and challenge identity norms through repetitive social interactions and behaviors conveying individual and collective knowledge, memories, and identities. This performative dimension doesn't merely reflect pre-existing identities; it plays a vital epistemological role. As we navigate our social world, it fosters self-discovery and a deeper understanding of ourselves and others. The repetitive, informative nature of these performances refines our self-understanding and deepens our grasp of identity dynamics in various social, cultural, and historical contexts. On the other hand, the concept of performativity is a fundamental mechanism in the discussion of gender identity and gender roles (Butler, 2006). It goes beyond viewing gender as a static biological imperative, assuming instead that gender is a dynamically and socially constructed category. It is perpetually shaped and reshaped by our individual and collective actions, experiences, and interpretations. Our everyday behaviours and interactions—ranging from how we dress and speak to our various behavioural patterns—serve as enactments of our gender roles (Voto, 2022). These actions and behaviours constitute performances of gender, each one subtly informing and shaping our gender identities. These performances are not merely descriptive or representative; they are constitutive. They don't just express or signify an already existing gender identity but participate in the process that creates and modifies our gender identities. Therefore, performativity in gender is not a process of revealing an identity but a dynamic process of creation and transformation. Every performance, has a tangible impact on our identities, contributing to an ongoing evolution of our sense of self.

2.2.2. When recognition fails: understanding the consequences of misidentification in human and artificial perception

The concept of misgendering has been thoroughly explored in the previous sections. Given the previous explanations, the following will briefly recapitulate the primary notions and add further points of discussion. For the purpose of this study, the concept of misidentification has been closely linked to the one of misgendering. While the first comprises a broader definition, the latter strictly concerns the act of referring to someone using words that do not correctly reflect the gender with which they identify, causing distress to the person involved (Dembroff & Wodak, 2018; Keyes, 2018; Sue, 2010). Misidentification, characterized as the inaccurate recognition of an individual or object, constitutes a phenomenon ubiquitously observed in both human cognition and artificial intelligence. The consideration of misidentification, alongside the acknowledgment of the resultant error and its consequences, presents an opportunity to mitigate the divide between human cognitive processes and artificial mechanisms. The ensuing discourse subsequent to such interactions is entwined with pre-existing norms, functioning as interpretative frameworks not directly tethered to the specific accuracy or inaccuracy of the identification of an individual or object. These frameworks underscore the intersections where diverse thought processes converge and exchange information, thereby emphasizing that our comprehension of identity and susceptibility to

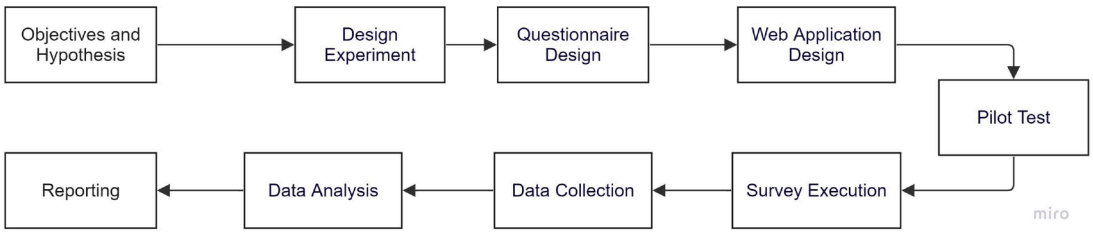


Fig. 1. Flowchart depicting the experimental design and methodology of the study.

misidentification are contingent upon the contextual intricacies of relationships and interactions, as well as the intrinsic characteristics of the entities being recognized. In this context, the etymol-ogy of the term *recognition* serves as a poignant reminder that the act of recognizing someone or something is an interpretative operation contingent upon the context in which the object, whether organic or inorganic, to be recognized has been previously encountered, known, or observed. Recognition, therefore, necessitates the retrieval from memory of experiences associated with identifying interpretative patterns, facilitating the association of something already known with information about to be acquired. In this vein, incorrect identification serves as a meaningful crucible for contemplating the reconsideration, refinement, and potential correction of interpretative patterns. A reciprocal relationship exists in this context, particularly within an intersubjective framework, between the subject and recognition. It can be asserted that the concept of the subject inherently incorporates the idea of the subject's recognition by others, signifying the emergence of the subject within intersubjectivity. This perspective enables the interpretation of various issues undoubtedly entwined with the notion of the subject, both philosophically and ethically-politically. The emphasis is on situating intersubjective recognition practices at the crux of the matter, underscoring that the recognition of subjectivity occurs exclusively among subjects; hence, it unequivocally warrants the designation of intersubjective. The existence of subjects constitutes both its presupposition and its outcome.

2.2.3. *Surviving the mainstream: the tactics of disidentification*

While disidentification's conceptual complexity is undeniable, its potential importance for advancing fairness in AGR systems must not be overlooked.

Understanding it could be a pivotal tool for guiding future developments in artificial intelligence.

Disidentification, as outlined by José Esteban Muñoz in his work *Disidentification: Queers of Color and the Performance of Politics*, refers to a survival strategy used by minority identities (Muñoz, 1999). This strategy employs tactics from prevailing power structures to navigate a majority-dominated public sphere, sparking unique social relations and creating new public spaces for alternative identities.

To illustrate, consider the example of internet personality Charlieshe, who uses disidentification through humor to challenge conventional notions of acceptable bodies on major online platforms (Bridges, 2021). Muñoz highlights how humor can transform sustained anger into a call for activism, reclaiming societal space dominated by norms. Another example is *RuPaul's Drag Race*, a platform where queer individuals re-contextualize and reclaim dominant narratives, challenging norms, shaping their identities, and empowering themselves.

Disidentification plays a crucial role in various queer theoretical frameworks, aiming to establish acceptance and visibility in public spaces. These spaces paradoxically offer opportunities for expressing and negotiating identities while being heavily influenced by societal norms that can hinder certain identities from recognition and acceptance.

This dual dynamic underscores disidentification's complexity, reflecting the struggle for representation while acknowledging systemic barriers. Despite its theoretical significance, practical limitations

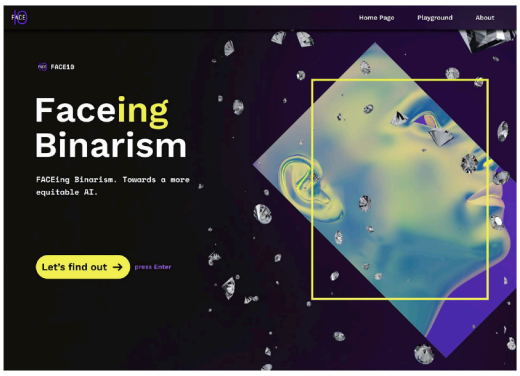


Fig. 2. Home-page of the web-application.

prevented its integration into our survey due to challenges in capturing its essence visually or through binary questions within our timeframe.

3. FACEing binarism

In light of the comprehensive discourse previously delineated regarding the concepts of identification, misidentification and disidentification, we designed and structured the online survey with the intent of reflecting these theoretical constructs. This survey was aimed at operationalizing these concepts, thereby enabling an analysis of their potential manifestations in real-world contexts.

The first Fig. 1 shows the pipeline of the experimental procedure.

Following the development of the hypothesis and the conceptual design of the experiment, we proceeded to construct the survey. We employed a mixture of closed-ended and Liker scale questions based on the type of responses needed for the final analysis. Crucially, the survey was embedded within a web application (Fig. 2).

The application was designed not only to host the survey, but to create an interactive experience enhancing responsive engagement. A unique feature of the survey design is the incorporation of a second web page activated within the *misidentification* section of the main survey, leveraging an Automatic Gender Recognition (AGR). This design choice aims to dig into the realm of human-computer interaction, with the ultimate purpose of exploring how individuals respond when artificial intelligence systems challenge their gender identity. The participants' reaction has been analyzed trough the second part of the survey.

The forthcoming paragraphs will delineate the survey's structural division, focusing on the salient topics that pertain to identification, misidentification and disidentification. Additionally, it will present an overview of the general statistics derived from the survey responses, providing an initial insight into the landscape shaped by these complex phenomena.

3.1. *Survey structure*

The complexity of gender identity, spanning binary to non-binary expressions, poses challenges for automated facial recognition technologies. Cultural norms add nuance to the intra-action between identity

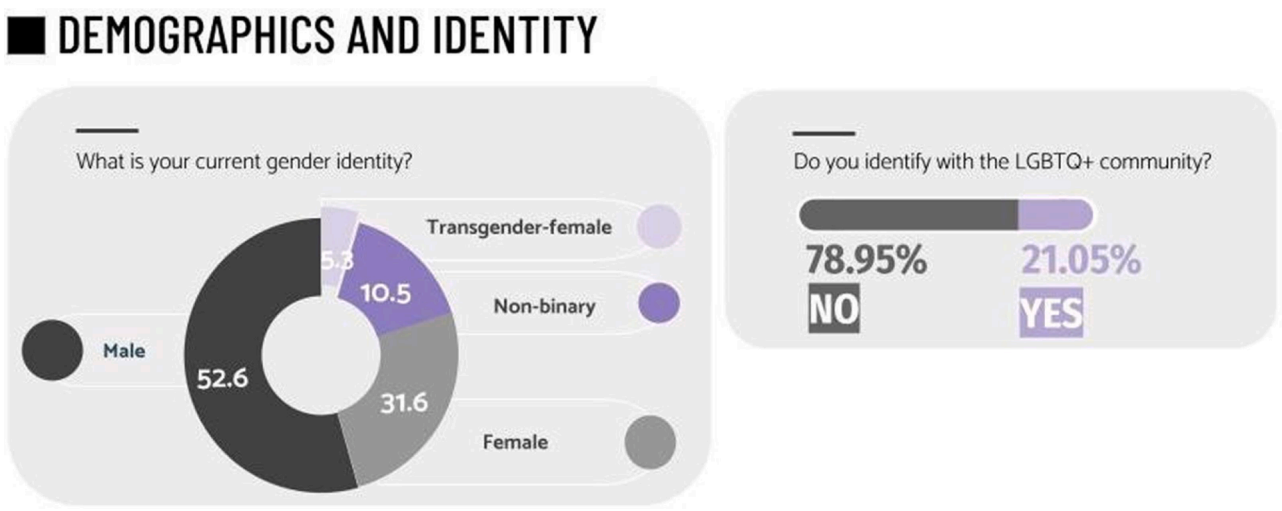


Fig. 3. Graphical representation of gender identity distribution among survey participants.

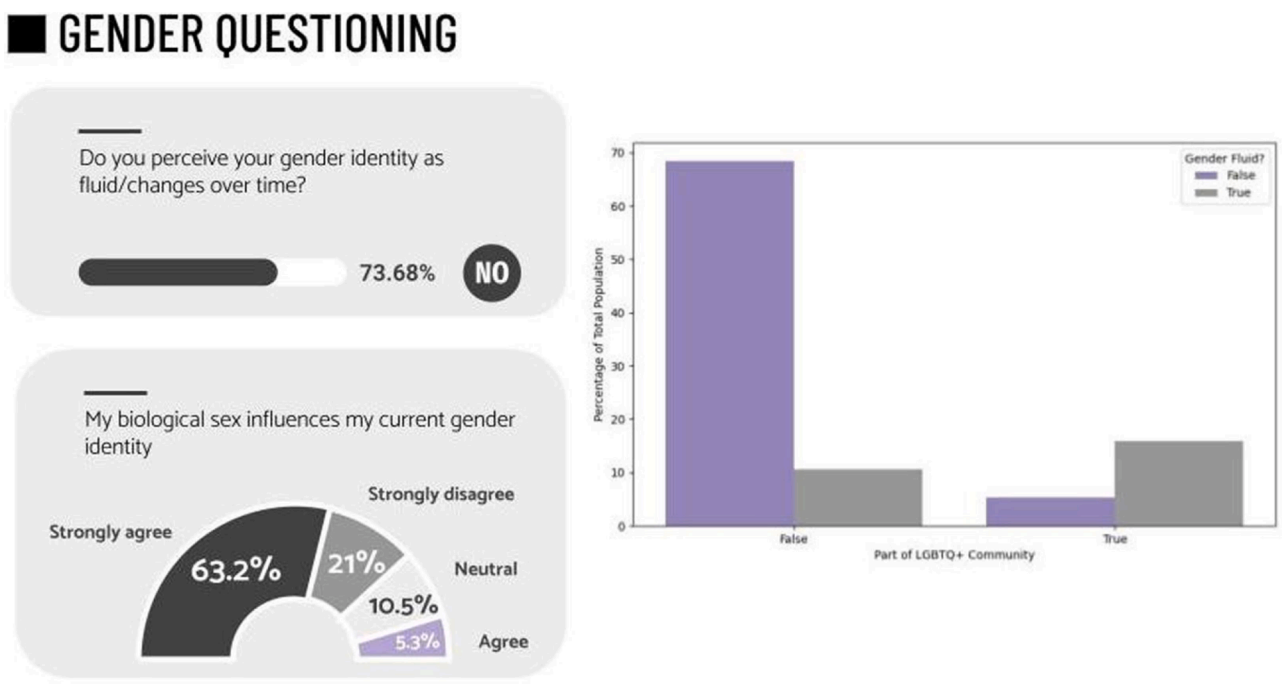


Fig. 4. Graphical representation of Gender Questioning segment.

and technology.

Our methodology aims to understand these dynamics, laying the foundation for comprehension between technology and diverse gender identities, ensuring resonance with both technical and societal aspects.

3.1.1. Gender and demographics

The first section of our survey was designed to accumulate pivotal demographic data pertinent to our research objectives and potentially beneficial for future comparative analyses (Fig. 3). This segment predominantly focuses on gathering information regarding the participants' current gender identity, an aspect central to understanding the nuances of identification processes.

3.1.2. Gender questioning

The second segment of the survey forms the initial part of the identification-focused section (Fig. 4). It is designed to delve into the intricacies of gender identity in relation to biological sex and its perceived fluidity. The section begins by inquiring whether the participants' biological sex influences their current gender identity. Following this, the survey probes into the concept of gender fluidity, asking participants if they perceive their gender identity as being dynamic and changeable over time. This exploration is further nuanced by examining its appearance in relation to the participants' membership to the LGBTQ+ community. This section, as reflected by the term 'Gender Questioning', aspires to reflect the investigative approach towards understanding how individuals perceive and construct their gender identities, especially in the context of the performative and fluid nature of gender. This aligns with the emphasis on identification as a dynamic process of assimilating and transforming self-identity based on interactions with others and the influence of societal norms and expectations.

3.1.3. Identity and performativity

Continuing our examination within the Identification section, we

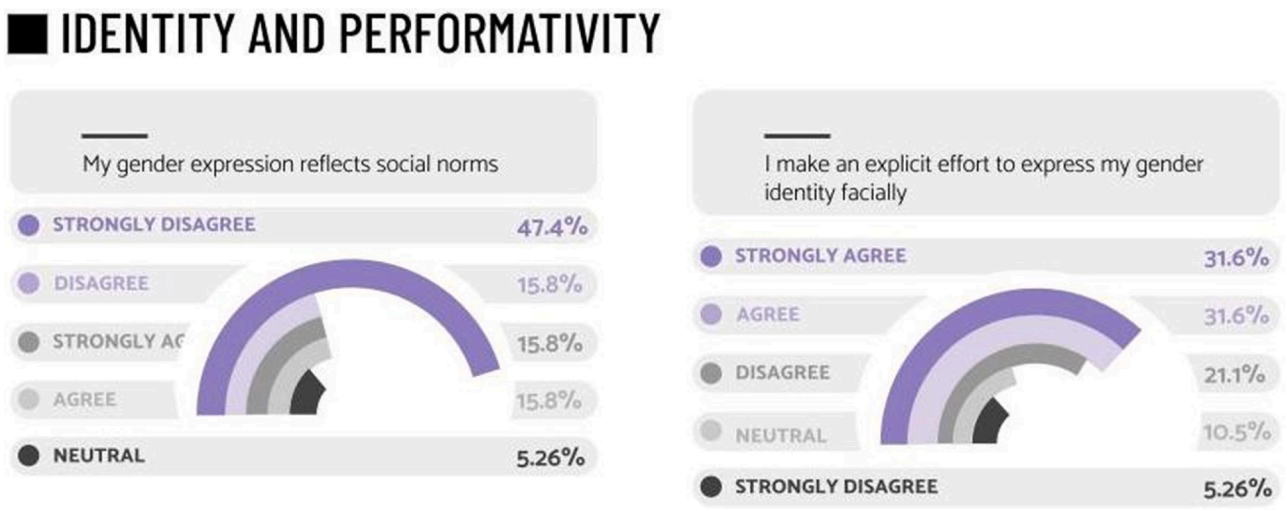


Fig. 5. Graphical representation of Identity and Performativity segment.

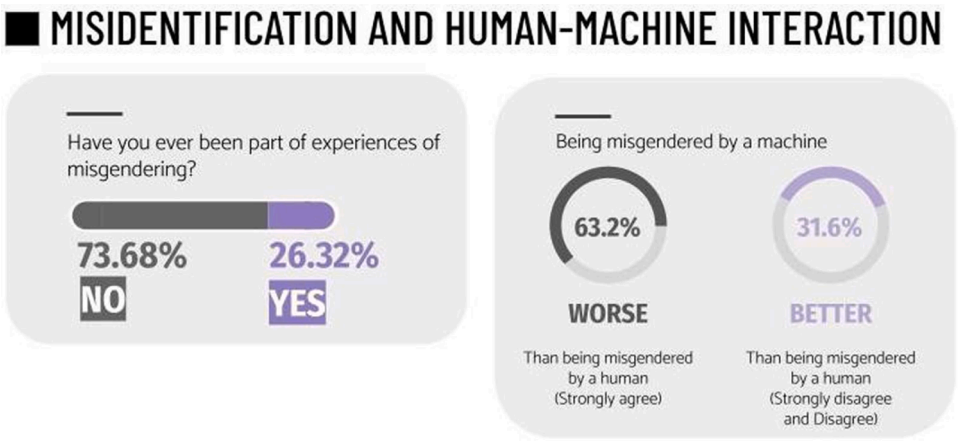


Fig. 6. Graphical representation of machine versus human misgendering experiences based on survey data.

delve deeper into the realm of identity and its performative aspects (Fig. 5). This section aims to scrutinize the following two questions of interest. The first segment explores participants' perceptions regarding the alignment of their gender expression with prevailing societal norms. It seeks to ascertain whether individuals perceive their gender expression as naturally conforming to societal expectations. The second segment delves into the area of facial expressions as a deliberate means of explicitly conveying one's gender identity. This query seeks to gauge whether participants make conscious efforts to utilize facial expressions in the communication of their gender identity. Within the overarching theme of 'Identity and Performativity', our goal is to gain insights into the complex dynamics of identity construction and self-expression, particularly in relation to societal norms and the deliberate choices individuals make in expressing their identities.

3.1.4. *Misidentification and human-machine interaction*

Within the domain of misidentification, our survey segment focuses on the interaction between misgendering experiences and evolving landscape of human-machine interfaces (Fig. 6). Firstly, we inquire about instances of individuals encountering misgendering by fellow humans. This exploration seeks to understand the frequency and nature of such experiences. Subsequently, we assess the level of discomfort individuals have encountered during these misgendering encounters, both in human interactions and within a simulated context. This allows us to gauge the emotional impact of misgendering across different scenarios. Moreover, we extend our investigation to consider the comparison between misgendering experiences involving humans and those involving machines. Participants are asked to share their perspectives on whether being misgendered by a machine is perceived as more distressing or less distressing than similar experiences with humans. It's essential to note that the disidentification aspect, while relevant, was not directly examined in our survey. The rationale behind this omission lies in the need for a distinct analytical approach, which would require specific training of participants on the subject.[1] Consequently, our hypotheses regarding disidentification remain unexplored terrain, awaiting further investigation in subsequent analyses.

3.2. *Preliminary findings*

Our study uncovered several preliminary findings that shed light on the complex relationship between gender identity, gender expression, and experiences of misgendering. It's important to emphasize that these findings are initial observations and require further investigation for a comprehensive understanding.

One notable finding is the positive correlation between identifying as

[1] Computer vision dazzle can be considered as an already attested training example being a form of facial makeup specifically crafted to conceal facial features and impede AFA

nonbinary and disagreeing with the idea that one's current gender identity is influenced by biological sex. This suggests that individuals who identify as non-binary tend to reject the notion that their gender identity is determined by their biological characteristics.

Additionally, our analysis revealed a positive relationship between the per-ceived influence of biological sex on current gender identity and the alignment of gender expression with social norms. Lower values in the latter variable corresponded to lower values in the former, indicating that participants who believed their gender identity was influenced by their biological sex also tended to express their gender in ways that conform to societal expectations.

Another interesting observation was the strong correlation between perceiving gender as fluid and having experienced previous instances of misgendering. This finding suggests a connection between considering gender as fluid and being more likely to have encountered misgendering in the past.

Furthermore, we identified a negative correlation between the statement *I make an explicit effort to express my gender identity facially* and the level of discomfort experienced after intentional misgendering. Participants who disagreed with this statement, indicating that they put less effort into facially expressing their gender identity, tended to report higher levels of discomfort following instances of misgendering.

In addition, we found a consistent pattern where participants who believed that being misgendered by a machine is worse than being misgendered by a human also tended to believe that being misgendered by a machine is better than being misgendered by a human. This suggests that participants were generally consistent in their views on the relative severity of machine and human misgendering.

Finally, our examination of participants' opinions on the comparison between being misgendered by a machine and being misgendered by a human indicated that a substantial number of individuals leaned towards the idea that being misgendered by a machine is worse, as evidenced by a higher number of participants choosing *Strongly Agree* with this statement. Interestingly, there appeared to be less differentiation between those who believed it was better and those who perceived no significant difference between machine and human misgendering.

It's important to reiterate that these findings are preliminary, and further research is needed to delve deeper into these relationships and provide a more nuanced understanding of the dynamics at play.

3.3. Research questions: core insights

3.3.1. RQ1: To what extent does an individual's gender identity align with societal norms, and how is this reflected in their appearance, particularly their facial expression?

Individuals who identify as non- binary often disagree with the notion that their gender identity is influenced by biological sex. This suggests that those adhering to traditional gender binaries, such as male and female, are more likely to believe in a link between biological sex and gender identity. Moreover, there appears to be an association between individuals perceiving their gender identity as influenced by biological sex and their belief that their gender expression conforms to societal norms. This could highlight societal pressures or expectations to conform, which may not be felt by those who do not see their biological sex as influential in their gender identity. Another insight is that those who make a conscious effort to express their gender identity through their facial expression might be more prone to the emotional distress caused by intentional misgendering, emphasizing the need for societal acceptance and recognition.

3.3.2. RQ2: Is there a perceived difference between being misgendered by a machine and being misgendered by a human?

There is a notable association between perceiving gender as fluid and prior experiences of misgendering. This indicates that individuals with fluid gender identities may be more susceptible to misunderstandings, whether societal or technological. A key insight is that a significant number of participants view machine misgendering as more problematic than misgendering by humans, possibly due to expectations of higher accuracy from machines or the impersonal nature of machine interactions making misgendering more distressing. Conversely, an almost equal number of responses suggest that machine misgendering is either less concerning or similar to human misgendering, pointing to a need for further investigation into these varied perceptions.

4. Ethical considerations: implications for gender identity and human–AI interaction

The upcoming advancement of Automated Facial Analysis (AFA) technologies, offers promising opportunities in various sectors. They are increasingly used as programmed entities to achieve specific outcomes, such as detecting, identifying, or classifying someone based on a facial image. Yet, it raises substantial ethical concerns, especially as these technologies intersect with deeply personal facets of human identity, like gender.

Their widespread use prompts inquiries regarding their influence on the perception and expression of identity. This is why is necessary to highlight the importance of monitoring and managing these technologies and promoting a deeper and more supportive alliance between humans and non-human agents when they interact in the perception of humanness.

Within this perspective, in the complex relationship between the features of automated recognition systems and the experience of being recognized, two fundamental ethical issues are central: first and foremost, gender classification in facial recognition technology can reinforce gender norms and stereotypes; secondly, machine recognition when responses to human identity traits can influence the socio-cultural construction of gender categories.

We will now discuss some ethical challenges and suggest future research directions in this field.

4.1. Identification, ethical accountability and the risk of embedded biases

The ethical accountability of AFAs in this domain is not merely about technological accuracy. It extends to the recognition and respect of the myriad ways in which gender identity is expressed and experienced. AFA systems that fail to acknowledge this diversity can inadvertently enforce rigid, binary notions of gender, leading to the exclusion or misrepresentation of non-binary and transgender individuals.

4.2. Visual data and gender predictions

The coherence of gender predictions made by AFAs models is increasingly questioned. AFAs predictions of 'gender' might be an assemblage of unrelated visual cues, far removed from the social understanding of gender. This insight is particularly relevant in our discussion on identification and technology. Recent studies have shown that gender expression can be discerned from visual datasets (Mane & Shah, 2019), even when reduced to basic elements like average color values. This finding underscores a critical ethical concern: the perpetuation of stereotypes and biases in AI systems. The removal of gender artifacts from datasets, intended to mitigate bias, can paradoxically be counterproductive. It often eliminates crucial information necessary for tasks like object or scene recognition. This phenomenon underlines the necessity of inclusive design, demonstrating that biases are not merely a matter of representation but are deeply embedded in the datasets that feed these systems.

4.3. Fairness-through-awareness

The shift from the 'fairness-through-blindness' approach in AI, which attempts to mitigate bias by erasing gender indicators from data, to a 'fairness-through-awareness' approach is crucial, acknowledging

the complexity of gender as a social construct and advocating for the adaptation of algorithms to recognize and respect gender differences. Furthermore, in the construction of visual datasets, we advocate for a participatory approach where types of gender artifacts are included or excluded based on contextual decisions. This is particularly effective when gender data is self-reported, allowing for a more accurate and representative dataset. Such a participatory approach not only aligns with our focus on a diverse spectrum of gender identities but also is essential for AI practitioners to discern between gender artifacts that might perpetuate harmful stereotypes and those that represent important social distinctions.

The relationship between technology and gender identity poses ethical challenges that require in-depth analysis. Interdisciplinary research in gender studies, computer science, and ethics is essential to addressing these challenges and promoting a more responsible and equitable use of technology. Additionally, it is necessary to consider the epistemic value of facial images, and thus facial data, and the effect of their reproduction in contemporary culture. Facial images in human communication are a fundamental artifact that has not yet been adequately analyzed in terms of their influence on identity representation. Understanding the interaction between facial images, facial data and facial recognition systems can provide further valuable insights about the impact of the perception of individual and collective identity within our society.

5. Concluding remarks and future directions

Reflecting on identification processes, whether stemming from human or artificial agency, brings forth an understanding of subjectivity fundamentally shaped by intersubjectivity—an element actively participating in and emerging from the collective human experience. Indeed, within this intersubjective dimension, identity performances intersect with a wide spectrum of identification processes, encompassing the establishment of normative identities to the engagement in resistance practices, exemplified in cases of disidentification. In every instance, intersubjectivity holds primacy over individual entities and, consequently, over the assertions, rights, duties, and responsibilities of human individuals. This doesn't involve reducing the individual subject to a collective super-subject, nor does it tether their identity to a group identity. Instead, grounding the subject in intersubjectivity represents, perhaps, the sole maneuver through which one can envision more sustainable alliances with technologies.

This research offers a comprehensive examination of the challenges and biases present in Automatic Gender Recognition (AGR) systems, particularly their tendency to categorize gender identities within a restrictive binary frame-work. Our findings reveal that such systems, while technologically advanced, often fail to capture the complex, fluid nature of gender identity, resulting in the marginalization and misgendering of individuals who do not conform to traditional gender binaries. This study contributes significantly to the field by employing an innovative approach, utilizing an online survey in conjunction with an AGR system to simulate and analyze experiences of misgendering. Through this methodology, we were able to highlight the stark differences in how machines and humans perceive and categorize gender, drawing attention to the inherent biases and limitations of current AI technologies in this area. Our research underscores the importance of moving beyond purely technical solutions and adopting a holistic approach that considers the ethical, societal, and cultural dimensions that are integral to the development and deployment of facial recognition technologies. Importantly, the study also emphasizes the critical role of facial images, not merely as identifiers but as powerful symbols embedded with societal narratives and representations of identity. By exploring the performative aspects of gender and the implications of misgendering by automated systems, we shed light on often-overlooked facets of identity in technological contexts. Looking ahead, this initial study lays the groundwork for future research, setting the stage for further exploration into the recognition of identity and otherness in the age of machines. Our ongoing research aims to refine and expand our understanding of these issues, contributing to the broader academic dialogue and fostering the development of more inclusive, ethical, and responsible AI technologies.

CRediT authorship contribution statement

Elena Beretta: Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cristina Voto:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Elena Rozera:** Writing – original draft, Visualization, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Appendices

The survey

Welcome to our research study! We are interested in understanding how machine-classified and self-perceived gender identity are conveyed in Facial Recognition Technologies. You will be presented with information relevant to this issue. Then, you will be asked to answer some questions about it. Your responses will be kept completely confidential.
The study will take you around 10 min to complete. Your participation in this research is voluntary, and you have the right to withdraw at any point during the process.
By clicking the button below, you acknowledge the previous information:
I consent, begin the study.
I do not consent, I do not want to participate.
Demographic Questions
What is your current gender identity?
Male
Female
Transgender female
Transgender male
Non-binary
Agender/Don't identify with any gender
Prefer not to answer
Gender not listed. Prefer to self-describe below:

(*continued on next page*)

(*continued*)

2. What is your age?
Do you identify with the LGBTQ+ community?
No
Yes
Identification
Please select how much you agree or disagree with the following statement.
My biological sex influences my current gender identity.
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree
Do you perceive your gender identity as fluid/changes over time?
No
Yes
I make an explicit effort to express my gender identity facially.
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree
My gender expression reflects social norms.
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree
Misidentification
Have you ever been part of experiences of misgendering?
No
Yes
9. How much discomfort did you experience?
Rate your experience on a scale from 1 to 10, 1 being 'No Discomfort' and 10 being 'Extreme level of Discomfort'
1 2 3 4 5 6 7 8 9 10
Did the machine guess the right gender?
No
Yes
Have you ever found yourself in a similar situation?
No
Yes
12. How much discomfort did you experience during the simulation? Rate your experience on a scale from 1 to 10, 1 being 'No Discomfort' and 10 being 'Extreme level of Discomfort'
1 2 3 4 5 6 7 8 9 10
Being misgendered by the machine is worse than being misgendered by a human.
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree
Being misgendered by a machine is better than being misgendered by a human.
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree
There is no difference between being misgendered by a machine and being misgendered by a human
Strongly Agree
Agree
Undecided/Neutral
Disagree
Strongly Disagree

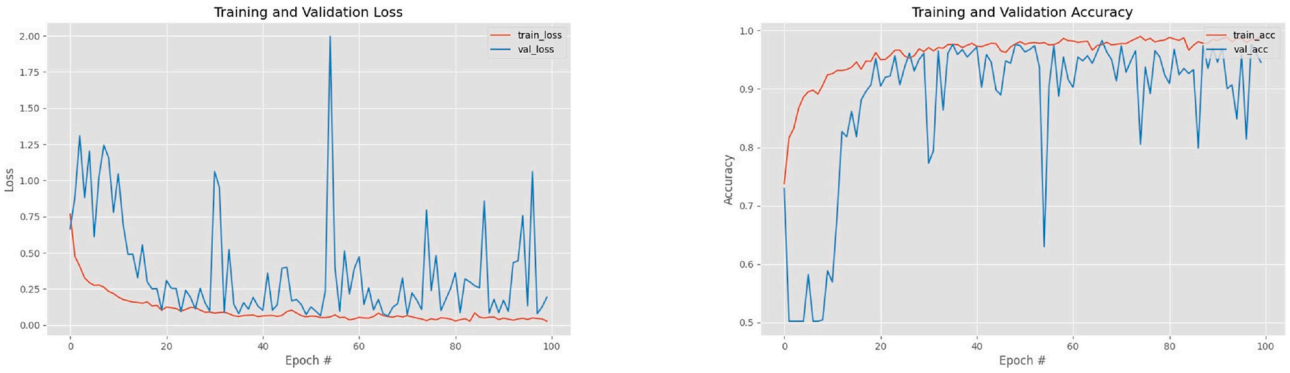
The AGR system

The system operates on two primary stages: the first being the detection and extraction of the face from a real-time image and the second involving the classification of the extracted face into one of the two gender categories: man or woman. The designed AGR model has been trained using Keras, and it employs a multi-layered CNN for the purpose of gender classification. In addition to its core functionality, the AGR system has been integrated in a Flask application to create an interactive experience for the participants.

The web application allows the users to submit facial images via an intuitive user interface and receive a gender prediction in return.

Data collection and preprocessing

The CNN model in this study was trained using a subset sourced from the CelebFaces Attributes (CelebA) dataset, which contains over 200,000 annotated celebrity images. For our research, 2000 images were selected, with an even split between *men* and *women* labels. These images underwent preprocessing to ensure proper cropping and alignment.



(a) Training and Validation Loss (b) Training and Validation Accuracy

Fig. 7. Training results showing loss and accuracy plots.

Architectural design of CNN

The CNN, tailored for gender detection in facial images, was designed for optimal performance without delving into a broader optimization of gender detection systems. It accepts input images of 96 × 96 × 3 dimensions (height, width, RGB channels). Following the input is a 3 × 3 convolutional layer paired with a Rectified Linear Unit (ReLU) activation. Batch normalization accelerates learning post-convolution, succeeded by a max pooling layer to retain significant features while reducing spatial dimensions. Dropout layers are incorporated post-pooling to counter overfitting, operating at 25 % initially and 50 % later. The model then flattens the data for the dense layer with 1024 neurons, followed by another ReLU and batch normalization. Conclusively, an output layer with a sigmoid activation function is used for binary classification, predicting the gender based on the input image.

Feature extraction

The CNN automatically learns a robust set of hierarchical features directly from the image data. Initial layers discern basic attributes such as edges and colors. In deeper layers, these features amalgamate to discern intricate patterns and facial attributes, like eye or nose shape, potentially indicative of gender.

Training, testing and optimization

The model is trained on 80 % of the dataset, with images preprocessed to 96 × 96 × 3 dimensions and normalized to a [0,1] pixel intensity range. To optimize the model's generalization to unseen data, the Adam optimizer was chosen for its efficiency and low memory footprint, set with a learning rate of 1e-3 that decays over epochs to ensure balanced convergence. Dropout layers are incorporated to mitigate overfitting by randomly nullifying input units during training updates. In tandem, an ImageDataGenerator facilitates data augmentation, enhancing the training set by applying diverse image transformations, thus fostering better model generalization. Training proceeds over 100 epochs using a batch size of 64, with performance evaluated on a validation set after each epoch. Once trained, the model is assessed on the remaining 20 % of the dataset, predicting gender probabilities benchmarked against true labels.

Evaluation metrics

The gender detection model's performance was evaluated using various metrics to assess its predictive abilities. Accuracy, the ratio of correct predictions (true positives and negatives) to the total, served as the primary metric due to our binary classification task. We also monitored binary cross-entropy loss, appropriate for binary classifications, to measure the disparity between predicted probabilities and actual labels. The training goal is to minimize this loss, ensuring predictions align with true labels. In the plots 7a 7b, tracking loss and accuracy against epochs, the model's consistent performance on the training set is evident, marked by low loss and high accuracy. Contrastingly, the validation metrics exhibit fluctuations. This divergence, especially the unstable validation accuracy in comparison to the stable training metrics, hints at potential overfitting. Essentially, the model adeptly captures training data nuances but may falter in applying this understanding to unseen validation data (Fig. 7).

Optimization and tuning

The optimization and tuning of the CNN model directly influences the model's ability to generalize and perform accurately on unseen data. In the model designed, the Adam optimizer was chosen due to its efficiency and lower memory requirements. The learning rate, determining the step size at each iteration while moving toward a minimum of the loss function, was set to 1e3, providing a balanced rate of convergence. Further, to prevent overfitting, Dropout layers were used in the architecture. This regularization technique randomly sets a fraction of input units to 0 at each update at training time, preventing overfitting. As part of the tuning process, an ImageDataGenerator was used for data augmentation, as mentioned above in the training paragraph.

References

Andrejevic, M., & Selwyn, N. (2020). Facial recognition technology in schools: Critical questions and concerns. *Learning, Media and Technology, 45*, 115–128. https://doi.org/10.1080/17439884.2020.1686014. URL: https://www.tandfonline.com/doi/full/10.1080/17439884.2020.1686014.

Balla, P. B., & Jadhao, K. T. (2018). IoT based facial recognition security system. In *2018 International Conference on Smart City and Emerging Technology (ICSCET)* (pp. 1–4). Mumbai, India: IEEE Xplore. https://doi.org/10.1109/ICSCET.2018.8537344.

Barr, A., 2015. Google mistakenly tags black people as "gorillas," showing limits of algorithms. Retrieved from https://www.wsj.com/articles/BL-DGB-42522.

Beauvoir, S. D. (1948). *The second sex*. New York: Alfred Knopf.

Belting, H. (2017). *Face and mask: A double history*. Princeton, NJ: Princeton University Press.

Bisogni, C., Castiglione, A., Hossain, S., Narducci, F., & Umer, S. (2022). Impact of deep learning approaches on facial expression recognition in healthcare industries. *IEEE Transactions on Industrial Informatics, 18*, 5619–5627. https://doi.org/10.1109/TII.2022.3141400. conference Name: IEEE Transactions on Industrial Informatics.

Bradford, B., Yesberg, J. A., Jackson, J., & Dawson, P. (2020). Live facial recognition: Trust and legitimacy as predictors of public sup-port for police use of new technology. *The British Journal of Criminology, 60*, 1502–1522. https://doi.org/10.1093/bjc/azaa032. URL:.

Bridges, L. E. (2021). Digital failure: Unbecoming the "good" data subject through entropic, fugitive, and queer data. *Big Data & Society, 8*. https://doi.org/10.1177/2053951720977882

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 1–15.

Butler, J. (2002). *Gender trouble*. routledge.

Butler, J. (2004). *Precarious life: The powers of mourning and violence*. verso.

Butler, J. (2006). *Gender trouble: Feminism and the subversion of identity*. London: Routledge, New York.

Christopher Hlongwane, T., Mathonsi, T. E., Du Plessis, D., & Muchenje, T. (2021). A review paper on facial recognition techniques in e-business. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 1702–1707). https://doi.org/10.1109/CSCI54926.2021.00323

Crenshaw, K. (2017). *On intersectionality: Essential writings*. The New Press.

De Lauretis, T. (1987). *Technologies of gender: Essays on Theory, film, and fiction*. Indiana University Press. URL: https://www.jstor.org/stable/j.ctt16gzmbr.

Dembroff, R., & Wodak, D. (2018). He/she/they/ze. *Ergo: An Open Access Journal of Philosophy, 5*. https://doi.org/10.3998/ergo.12405314.0005.014

Fausto-Sterling, A., & Stein, E. (2004). Sexing the body: Gender politics and the construction of sexuality. *Hypatia, 19*, 203–208. https://doi.org/10.1111/j.1527-2001.2004.tb01309.x

Halberstam, J. (2019). *Female masculinity*. Duke University Press.

Hall, S., & Du Gay, P. (1996). *Questions of cultural identity: Sage Publications*. Sage.

Hamidi, F., Scheuerman, M. K., & Branham, S. M. (2018). Gender recognition or gender reductionism? the social implica- tions of embedded gender recognition systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3173574.3173582. URL:.

Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). *Labeled faces in the wild: A database for studying face recognition in Unconstrained Environments*. Amherst: University of Massachusetts. Technical Report 07-49.

Karovaliya, M., Karedia, S., Oza, S., & Kalbande, D. R. (2015). Enhanced security for ATM machine with OTP and facial recognition features. *Procedia Computer Science, 45*, 390–396. https://doi.org/10.1016/j.procs.2015.03.166. URL: https://www.sciencedirect.com/science/article/pii/S1877050915004093.

Kaur, P., Krishan, K., Sharma, S. K., & Kanchan, T. (2020). Facialrecognition algorithms: A literature review. *Medicine, Science and the Law, 60*, 131–139. https://doi.org/10.1177/0025802419893168. URL:publisher: SAGE Publications.

Keyes, O. (2018). The misgendering machines: Trans/hci implications of automatic gender recognition. *Proceedings of the ACM on Human Computer Interaction, 2*. https://doi.org/10.1145/3274357. URL:.

Krithika, L. B., Venkatesh, K., Rathore, S., & Kumar, M. H. (2017). Facial recognition in education system. *IOP Conference Series: Materials Science and Engineering, 263*, Article 042021. https://doi.org/10.1088/1757-899x/263/4/042021. URL: publisher: IOP Publishing.

Kumar, V., Raghavendra, R., Namboodiri, A., & Busch, C. (2016). Robust transgender face recognition: Approach based on appearance and therapy factors. In *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)* (pp. 1–7). https://doi.org/10.1109/ISBA.2016.7477226. URL:.

Lacan, J. (1962). *L'identification*. M. Roussan, 2003.

Majumder, S., & Bhattacharya, S. C. (2021). A systematic literature review on the use of facial recognition in human resources using deep learning. *Interdisciplinary research in technology and management* (p. 5). Num Pages: CRC Press.

Mane, S., & Shah, G. (2019). Facial recognition, expression recognition, and gender identification. In V. E. Balas, N. Sharma, & A. Chakrabarti (Eds.), *Data management, analytics and Innovation* (pp. 275–290). Singapore: Springer. https://doi.org/10.1007/978-981-13-1402-521.

Meissner, W. W. (1970). Notesonidentification: I.originsinfreud. *The Psychoanalytic Quarterly, 39*, 563–589.

Merler, M., Ratha, N., Feris, R.S., Smith, J.R., 2019. Diversity in faces.

Mujtaba, D. F., & Mahapatra, N. R. (2019). Ethical considerations in AI based recruitment. In *2019 IEEE International Symposium on Technology and Society (ISTAS)* (pp. 1–7). IEEE. doi:10.1109/ISTAS48451.2019.8937920. iSSN: 2158-3412.

Muñoz, J. E. (1999). *Disidentifications: Queers of color and the performance of politics. volume 2*. U of Minnesota Press.

Preciado, P. B. (2013). *Testo junkie: Sex, drugs, and biopolitics in the pharmacopornographic era*. The Feminist Press at CUNY.

Ryu, H.J., Adam, H., Mitchell, M., 2017. Inclusivefacenet: Improving face attribute detection with race and gender diversity.

S, K.K., Vangara, K., King, M.C., Albiero, V., Bowyer, K.W., 2019. Characterizing the variability in face recognition accuracy relative to race. CoRR abs/1904.07325. URL: http://arxiv.org/abs/1904.07325, arXiv:1904.07325.

Scheuerman, M. K., Paul, J. M., & Brubaker, J. R. (2019). How computers see gender: An evaluation of gender classification in commercial facial analysis services. *Proceedings of the ACM on Human-Computer Interaction, 3*(144), 1–33. https://doi.org/10.1145/3359246. URL:.

Sekula, A., 1986. The body and the archive. October 39, 3–64. URL: https://www.jstor.org/stable/778312.

Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. John Wiley & Sons.

Tian, Y. L., Kanade, T., & Cohn, J. F. (2005). Facial expression analysis. In S. Z. Li, & A. K. Jain (Eds.), *Handbook of face recognition* (pp. 247–275). New York, NY: Springer. https://doi.org/10.1007/0-387-27257-712. URL:.

Voto, C. (2022). Artificialskinandbiopoliticalmasks, orhowtodealwithface—*habits. Signs and Society, 10*, 126–142.

Williford, J. R., May, B. B., & Byrne, J. (2020). Explainable face recognition. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI* (pp. 248–263). Springer International Publishing.

Wolf, K. (2015). Measuring facial expression of emotion. *Dialogues in Clinical Neuroscience, 17*, 457–462. https://doi.org/10.31887/DCNS.2015.17.4/kwolf. URL: publisher: Taylor & Francis eprint.

Wu, W., Protopapas, P., Yang, Z., & Michalatos, P. (2020). Gender classification and bias mitigation in facial images. In *Proceedings of the 12th ACM Conference on Web Science* (pp. 106–114).

Zeng, M., & Chiu, C. L. (2021). The use of facial recognition for online business with the perspective of customer adoption. In *2021 14th CMI International Conference Critical ICT Infrastructures and Platforms (CMI)* (pp. 1–7). https://doi.org/10.1109/CMI53512.2021.9663766